

Principles of Responsible Work

by James Bach, Jon Bach, and Michael Bolton

1. **Every non-trivial business comprises some set of services that enable it to function.** Examples include sales, accounting, R&D, customer support, etc. These services must be sufficiently reliable or else the business will collapse.
2. **Every service entails the risk of failure.** When failures occur, the business must be able to recognize them and recover. In regulated industries, risk management may be subject to specific process mandates.
3. **A “responsible person” is a natural person in a business who is reasonably competent, prepared, and accountable for some service that sustains or defines that business.** No matter what tools or processes are used within a business, someone must be responsible for them. To bear responsibility, a person must have sufficient capacity. For instance, neither a child nor a tool (such as AI) has the capacity (either legally or socially) to bear responsibility. Even adult humans may lack capacity, such as when an airline pilot has had insufficient sleep or is under the influence of drugs.
4. **A “responsible service” is one that is performed in good faith by a responsible person.** This may include interpreting and following procedures, improving skills, anticipating problems, and reporting to relevant authorities or clients, both inside and outside the business.
5. **Responsible services may incorporate any manner of tool, as long as the person performing that service can operate the tool safely and legally.** The effort and skill required to operate a tool safely increases with the complexity of the tool, the obscurity of its output, the amount of output produced per unit of time, and the baseline reliability of the tool when performing that task.
6. **Responsibility can be taken, shared, declined, or delegated, as long as there is a clear and reasonable protocol for doing so.** In the absence of such a protocol, the business is vulnerable to accusations of negligence. This is a principal topic of common law and the law of contracts, although specific laws and regulations may constrain how a business can distribute responsibility.
7. **Therefore, to avoid inefficiency, poor quality, and legal trouble, businesses must develop and maintain clear lines of responsibility, assure competence and readiness among responsible persons, put reliable tools in place, and maintain appropriate oversight of any delegated responsibility.**

Responsible Operation of AI

1. **AI cannot bear responsibility.**¹ AI is not a responsible person, and it would be meaningless to speak of a tool that operates in “good faith.” Therefore, it cannot provide a responsible service, nor can responsibility be delegated to it.
2. **An “AI agent” is always a tool operated by a natural person, irrespective of whether the person is monitoring it in real-time.**
3. **Thus, the operator of an AI agent always bears responsibility for its work.**^{2,3}
4. **The responsible operator cannot merely prompt and pray; they must assure adequate quality of the work.**

Therefore, the operator must...

5. **be sufficiently skilled in the use of the AI tool.**
6. **be sufficiently prepared to operate the tool in that context.**
7. **be sufficiently alert to risks, anomalies, or defects that may occur in the work**⁴.
8. **feel empowered (and actually have the power) to reject or remediate any work done by AI.** Otherwise, the operator becomes a scapegoat, a “moral crumple zone.”⁵
9. **avoid cognitive overload**⁶, **excessive cognitive debt**⁷, and **cognitive surrender**^{8,9}.

¹ See *Principles of Responsible Work*.

² There are relevant concepts in law that make employers liable even for *human* agents, such as *vicarious liability* and *apparent authority*. But, as a mere machine, AI does not have any societal standing and *cannot* be held liable for its actions.

³ Green, B. (2022). *The Flaws of Policies Requiring Human Oversight of Government Algorithms*. *Computer Law & Security Review*, 45, 105681. <https://doi.org/10.1016/j.clsr.2022.105681>

⁴ Klein, G., Pliske, R., Crandall, B., & Woods, D. D. (2005). *Problem Detection*. *Cognition, Technology & Work*, 7(1), 14–28. <https://doi.org/10.1007/s10111-004-0166-y>

⁵ Elish, M. C. (2019). *Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction*. *Engaging Science, Technology, and Society*, 5, 40–60. <https://doi.org/10.17351/ests2019.260>

⁶ Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2506.08872>

⁷ Ibid.

⁸ Shaw, S. D., & Nave, G. (2026). *Thinking—Fast, Slow, and Artificial: How AI is Reshaping Human Reasoning and the Rise of Cognitive Surrender*. PsyArXiv. https://doi.org/10.31234/osf.io/yk25n_v1

⁹ “So, You 10x’d Your Work”, <https://www.satisfice.com/blog/archives/488009>